

Utah State University

DigitalCommons@USU

All Graduate Theses and Dissertations

Graduate Studies

5-2012

The Reliability and Diagnostic Accuracy of the Yes/No Scapular Dyskinesis Test When Used By Graduate Assistant Athletic Trainers

Adam Raikes
Utah State University

Follow this and additional works at: <https://digitalcommons.usu.edu/etd>



Part of the [Medicine and Health Sciences Commons](#)

Recommended Citation

Raikes, Adam, "The Reliability and Diagnostic Accuracy of the Yes/No Scapular Dyskinesis Test When Used By Graduate Assistant Athletic Trainers" (2012). *All Graduate Theses and Dissertations*. 1203.
<https://digitalcommons.usu.edu/etd/1203>

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



THE RELIABILITY AND DIAGNOSTIC ACCURACY OF THE YES/NO SCAPULAR
DYSKINESIS TEST WHEN USED BY GRADUATE ASSISTANT ATHLETIC
TRAINERS

by

Adam Raikes

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Health and Human Movement

Approved:

Gerald Smith, Ph.D.
Major Professor

Dennis Dolny, Ph.D.
Committee Member

Trek Lyons, M.D.
Committee Member

Eadric Bressel, Ph.D.
Committee Member

Mark R. McLellan, Ph.D.
Vice President for Research and
Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, UT

2012

Copyright © Adam Raikes 2012

All Rights Reserved

ABSTRACT

The Reliability and Diagnostic Accuracy of the Yes/No Scapular Dyskinesia Test When
Used by Graduate Assistant Athletic Trainers

by

Adam Raikes, Master of Science

Utah State University, 2012

Major Professor: Dr. Gerald Smith

Department: Health, Physical Education, and Recreation

Context: Scapular motion evaluation is a necessary component of the upper extremity exam. Several methods exist, but most lack good reliability or diagnostic accuracy. The yes/no scapular dyskinesia test has the best of both measures but is untested on inexperienced clinicians. Objective: The purpose of this study was to evaluate the reliability and diagnostic accuracy of the yes/no scapular dyskinesia test when used by graduate assistant athletic trainers. Participants: The participants were college-aged students with no prior history of upper extremity fracture or nerve injury. Data Collection and Results: Participants were evaluated for scapular dyskinesia by a physician and 10 graduate assistant athletic trainers. Ratings were for normal or dyskinetic and then determination of side. Reliability was calculated using Gwet's AC1 statistic and diagnostic accuracy from standard 2x2 contingency tables. Results: Reliability was moderate ($AC1 = 0.48, p < 0.0025, 95\% \text{ CI } [0.147, 0.812]$) when side was not accounted for and moderate ($AC1 = 0.43, p < 0.0001, 95\% \text{ CI } [0.242, 0.632]$) when side-per-side

decisions were made. Sensitivity and negative predictive values were low to moderate (34.4%-66.2%, 8.9%-74.1%). Specificity and positive predictive values were moderate to high (50%-85.2%, 51.5%-95.2%). Accuracy was moderate (65.2%-69.4%) and positive and negative likelihood ratios were low (1.325-2.333, 0.675-0.769).

Conclusions: The reliability in this study was on par with previously published studies. Measures of diagnostic accuracy met or exceeded previous results. Clinically, to avoid false negative results and enhance the use of positive results, it appears necessary to combine methods and begin the evaluation with a gross assessment of whether or not dyskinesia is present and if it is to then evaluate which side is dyskinetic.

(75 pages)

PUBLIC ABSTRACT

The Reliability and Diagnostic Accuracy of the Yes/No Scapular Dyskinesia Test When Used by Graduate Assistant Athletic Trainers

The ability to accurately assess the motion of the scapulae is an important skill when evaluating injuries to the upper extremity. To date, several tests have been proposed and described as suitable methods for categorizing this motion. Scapular evaluations are challenging given the overlying musculature as well as the need to determine the relative timing of events. Previous tests' diagnostic accuracy has suffered as a result of these challenges.

Recently the Yes/No dyskinesia test has been proposed. This test eliminates much of the struggle with earlier evaluative methods by reducing the assessment to simply a yes or a no response as to the presence of abnormal scapular motion. This simplifies the evaluation and allows the clinician greater freedom to be unconstrained by multiple types of dyskinesia and the need to pigeon-hole a patient into a particular category.

The purpose of this study was to investigate the reliability and diagnostic accuracy of the test when used by relatively inexperienced clinicians. Seventeen Utah State University students had their scapular motion observed by the graduate assistant athletic training staff at Utah State University. The test demonstrated comparable reliability and diagnostic accuracy to previously published figures.

Adam Raikes

ACKNOWLEDGMENTS

I would like first to thank my committee members. I spent many hours in each of their offices, picking their brains, trying to find inventive ways of fully exploring this topic. Special thanks go to Dr. Smith for the countless articles, ideas, suggestions, and revisions that both this manuscript and the experiment itself went through. His work in each of these phases, as well as the step-by-step guidance through the process, allowed me to craft this project and to finally finish it. I am especially grateful to Dr. Bressel and Dr. Dolny off of whom I bounced numerous ideas and concepts and whose edits have been invaluable. Finally, Dr. Lyons devoted not only time and mental energy for the design of the project and manuscript but also his time to perform the evaluations on the participants. My committee made this project possible, so thank you.

I also want to thank my wife. She provided love and support that helped to get me through this entire process. She lent her ear so I could talk out the mathematics of the problem as well as heard me describe, ad nauseum, the motion and configuration of the scapulae. Her love, patience, ears, and support enabled me to figure out the minutae of this project and to persevere when it got frustrating.

Adam Raikes

CONTENTS

	Page
ABSTRACT.....	iii
PUBLIC ABSTRACT.....	v
ACKNOWLEDGMENTS	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER	
I. INTRODUCTION	1
II. LITERATURE REVIEW	4
Normal Scapular Motion.....	4
Abnormal Scapular Motion.....	10
Evaluation of the Dyskinetic Shoulder	12
Diagnostic Accuracy	21
Research Considerations	24
III. METHODS	26
Participants.....	26
Participant Recruitment	26
Participant Characteristics	27
Clinicians	28
Procedures.....	28
Training	28
SDT Test	29
Statistical Analysis	30
Reliability.....	30
Diagnostic Accuracy	31

IV.	RESULTS.....	33
V.	DISCUSSION.....	37
	Limitations	40
	Theoretical Applications	43
	Clinical Applications.....	44
	Future Research Considerations	45
	Conclusion	45
VI.	REFERENCES	47
VII.	APPENDICES	51
	Appendix A	52
	Appendix B	55
	Appendix C	57
	Appendix D.....	59

LIST OF TABLES

Table	Page
1. 2x2 Contingency Table for Evaluating Diagnostic Accuracy	21
2. Diagnostic Accuracy by Count	35
3. Diagnostic Accuracy	36

LIST OF FIGURES

Figure	Page
1. Scapular rotations.....	4
2. Number of participants engaging in select sports during high school or college	27

CHAPTER 1

INTRODUCTION

Shoulder pain and dysfunction are common complaints among athletes and active individuals. Common conditions include primary and secondary impingement syndromes, superior labrum anterior to posterior (SLAP) lesions and pre- SLAP insults to the labrum, partial and full thickness rotator cuff tears, and glenohumeral instabilities and laxities. Concomitant with these is a phenomenon referred to as scapular dyskinesis (Burkhart, Morgan, & Kibler, 2003; Ludewig & Cook, 2000). This dyskinesis reflects abnormal or inappropriate motion of the scapula as it articulates with the clavicle, the head of the humerus, and loosely the thoracic wall.

Numerous methods have been proposed in the literature for the quantitative and qualitative evaluation of scapular dyskinesis. These include Moiré topography (Warner, Micheli, Arslanian, Kennedy & Kennedy, 1992), the lateral scapular slide test (Odom, Taylor, Hurd, & Denegar, 2001), four category scapular dyskinesis test (SDT) (Kibler et al., 2002), and the yes/no dichotomous SDT (McClure, Tate, Kareha, Irwin, & Zlupko, 2009; Tate, McClure, Kareha, Irwin, & Barbe, 2009). Of these tests, the dichotomous method displays the most consistency in reliability (Tate et al., 2009) and in diagnostic utility as measured by sensitivity and specificity (Uhl, Kibler, Gecewich, & Tripp, 2009).

In the clinical setting, the ability to distinguish between normal and abnormal scapular motion is a key component of an upper extremity screen. Athletic training education programs include units on upper extremity assessment which often include scapular motion screening. However, instruction in scapular assessment is as varied

across programs as there are reported methods. Unlike other clinical diagnostic tools, such as the Lachmann's test for anterior cruciate ligament (ACL) ruptures, there is no gold standard of scapular tests.

To date, most of the research into the clinical evaluation of scapular dyskinesis has been in the hands of physical therapists and orthopedic surgeons, as well as biomechanists. There have been few, if any, in which the testers were athletic trainers and to date none that have been published where the testers were recent graduates. As Uhl et al. (2009) note, using experienced clinicians within the confines of a research endeavor provides the best case scenario for clinical diagnosis. While this is certainly true, it is impractical to measure the utility of a clinical test only by optimal conditions.

In the sports medicine realm, while a team physician or associated orthopedist may have experience with scapular evaluation, there will be very few instances where the level of experience under which Uhl et al. (2009) conducted their study could be met. The authors of that particular study are well-published in the area of scapular pathomechanics over the past fifteen years. What is likely is that athletes will be evaluated by an athletic trainer who may be a graduate assistant and thus a recent graduate and young professional with limited clinical experience or a staff athletic trainer who may or may not see scapular abnormalities frequently and therefore may have the requisite level of training and experience but may lack the specific experience in this area.

The first research purpose of this study is to determine whether graduate assistant athletic trainers with less than 2 years of clinical experience can use the dichotomous "yes/no" scapular dyskinesis test proposed by McClure and colleagues (2009) with

comparable reliability to that found by McClure et al. and Uhl et al. (2009). The second purpose is to assess the diagnostic accuracy of the test by use of evidence-based indices of the SDT when used by graduate assistant athletic trainers.

CHAPTER 2

LITERATURE REVIEW

Normal Scapular Function

Normal scapular motion occurs in three planes of motion. During the elevation phase of humeral forward flexion and coronal abduction, the scapula protracts anteriorly around the thoracic wall, the inferior angle rotates upward (or laterally), and the inferior angle posteriorly tilts away from the thoracic wall (see Figure 1). These motions maintain the bony configuration of the glenohumeral joint, maximize contact of the humeral head with the glenoid fossa, and assure sufficient clearance of the supraspinatus tendon and subacromial bursa in the subacromial space. These goals are achieved through the combined actions of the rotator cuff muscles – the supra- and infraspinatus, teres minor, and subscapularis – which secure the humeral head and the muscles of scapular motion and stabilization – the trapezius, serratus anterior, and both major and minor rhomboids (Kibler, 1998).

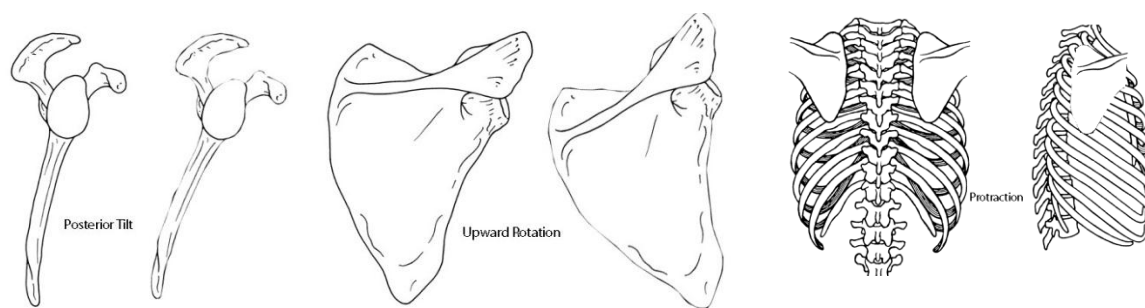


Figure 1. Scapular rotations (adapted from Cull, 1989).

During most overhead athletic and occupational endeavors, force is translated to an object or implement by the arm and hand. The transmitted force is largely the result of motion and force generation in the lower extremity and torso. The scapula acts as the functional link between the arm and the force producers of the trunk and legs (Burkhart et al., 2003; Kibler, 1998). To safely, accurately, and efficiently transmit this force to the arm requires the scapula to move against the thoracic wall and in conjunction with the humerus.

The scapular stabilizers work together to smoothly accomplish the three normal motions at the scapulothoracic joint. The muscles act in conjunction with one another and in simultaneous opposition to each other to create force couples. In normal motion, the upper and lower trapeziuses along with the rhomboids work opposite the serratus anterior to stabilize the scapula against the thorax (Kibler, 1998). To achieve acromial elevation, the lower trapezius and serratus oppose the upper trapezius and rhomboids to create posterior tilt and upward rotation of the scapula (Schmitt & Snyder-Mackler, 1999). If any of these muscles do not function optimally or at the same level of their force couple counterparts, the motion becomes dominated by the opposing factor and leads to malpositioning and a loss of smooth control of the motion (Kibler, 1998).

Scapulothoracic motion only describes half of the scapular role in shoulder motions. As described, the scapula moves to maintain the bony configuration of the glenohumeral joint as the ball-and-socket relationship translates through three dimensional space. To maximize stability, this motion must occur appropriately and simultaneously with the motion caused by the extrinsic shoulder musculature.

The tandem motion of the glenohumeral joint and the scapulothoracic joint has been termed scapulohumeral rhythm. Early measures of scapulohumeral rhythm demonstrated a ratio of 2:1 (scapula:humerus) through the total arc of humeral motion. Recent research has clarified this ratio considerably. McQuade and Smidt (1998) noted that early assessments of this rhythm were done with static arm positions via two dimensional radiographic films. The authors note the fallacy of this method of measurement is that the scapula moves in three dimensions. To constrain the motion to two dimensions as is done for X-rays negates one plane and changes the overall assessment of the joint's motion.

The advent of three-dimensional motion tracking has made it possible to improve upon the early methods of scapular motion measurement. McQuade and Smidt (1998) demonstrated that the scapulohumeral rhythm is a dynamic variable which changes with the relative position of the arm to 100% elevation and which alters depending on the presence of weighting or non-weighting. During unloaded (passive) arm motion, the authors found greater variability in the ratio, ranging from a 7.9:1 ratio at the initiation of motion to a 2.9:1 ratio during the final 20 degrees. Light loads showed less variability but did demonstrate an increasing, rather than decreasing rhythmic ratio during elevation, as did heavy loads.

Additionally, McQuade, Dawson, and Smidt (1998) demonstrated that repetitive motion sufficient to induce fatigue in the scapular stabilizer force couples resulted in a decrease in the scapulohumeral rhythm. Specifically, “fatigue tends to result in destabilization of the scapula or compensatory increased rotation primarily in the

midrange of arm elevation.” This has the impact of altering the normal kinematic relationship between the scapula and the humerus creating dyskinesis. The potential results are malpositioning of the humeral head in the ball-and-socket configuration and insufficient elevation of the acromial arch. Furthermore, the rotator cuff is placed at additional strain to stabilize the glenohumeral joint. Combined these potential results place the athlete at increased risk of debilitating shoulder injuries.

McQuade and Smidt (1998) and McQuade et al. (1998) presented several methodological challenges. The authors acknowledge the three-dimensionality of scapular motion and attempt to overcome this through 3D instrumentation (McQuade & Smidt, 1998). To do so, the authors placed surface markers over the deltoid tuberosity and over the acromion process. These positions were confirmed on nine separate subjects for each of nine positions of elevation. The use of a separate subject for each position was done to minimize risk from multiple X-ray exposure (McQuade & Smidt, 1998). However, it presumes that the anatomic make up of each individual's acromion and deltoid tuberosity was identical which is unlikely. Also, it relies on the use of two dimensional X-ray techniques to confirm the three dimensional positioning of the sensors over the various landmarks which were criticized in the literature review of the McQuade and Smidt article.

The McQuade et al. (1998) article posed an additional questionable method. EMG data were collected by the authors to confirm the onset and progression of fatigue in the upper and lower trapezii, serratus anterior, and middle deltoid. The middle deltoid was chosen as a comparison mark to show concomitant fatigue of the scapular stabilizers with

fatigue of the humeral elevator which reflects that the scapula destabilizes with fatigue at approximately the same time that the arm begins to rely on additional musculature to achieve elevation (p. 75). However, the rhomboids were left out of the stabilizer EMG data and these muscles reflect a significant part of the stabilization force couple. To ignore this muscle group may explain the compensatory motions because the work of the rhomboids to provide those stabilization actions is unknown.

Finally, both articles relied on only two sensors, one on the deltoid tuberosity and one on the acromion. While these positions reflect the relationship of these two landmarks, they do not demonstrate the motion occurring at the rest of the scapulothoracic joint. Protraction occurring in the initial stages of arm elevation may not be reflected sensitively enough at the acromion when related to the deltoid tuberosity and thus this model may not fully represent the true range of motion of the scapula during arm elevation. It does demonstrate the dynamic nature of the motion and that a linear relationship may be inappropriate for describing the coupled motions of the glenohumeral and scapulothoracic joints.

In a study of scapular position in cases of impingement, Lukasiewicz, McClure, Michener, Pratt, and Sennett (1999) overcame this particular shortcoming of the McQuade et al. (1998) study. To create a more accurate representation of the scapula's position, six landmarks were used: the spinous processes of C7 and T7, the medial scapular border at the root of the spine of the scapula, the inferior angle of the scapula, the posterior angle of the acromion, and the olecranon process. Measurements were taken at three static positions: rest, 90° of scapular plane elevation, and maximal elevation (p.

576). The authors found decreased posterior tilting and increased scapular elevation in the participants with impingement (p. 578).

The utilization of the six motion tracking points allows for a more accurate model of scapular position. The spinous processes allow for the analysis of medial-lateral and superior-inferior relative positions which the McQuade et al. (1998) study did not. The posterior acromion angle provides a reference point for determining tilt when compared with the inferior angle of the scapula. The olecranon process makes it possible to assess upward rotation.

The major limitation that Lukasiewicz et al. (1999) acknowledged is the use of static joint angles. Resting position makes sense as a static position. However, based on the findings in McQuade et al. (1998) regarding the dynamic nature of scapular motion across the total arc of humeral elevation, limiting findings to just two other points along the continuum negates much of the motion available. Thus while the three dimensional properties improve upon the McQuade et al. experiments, the loss of functional motion in testing represents a significant impediment to applicable results.

Lukasiewicz et al. (1999) further noted the difficulty in obtaining measurement points because of the difficulty palpating some of the landmarks. Though the study demonstrated intraclass correlation coefficients of 0.88 to 0.99 and a standard error of measurement of less than 2° , the consistency of results does not mean that the skin sensors accurately overlayed the actual landmarks, particularly in individuals on whom palpation was difficult. X-ray marking would have improved this study by conclusively denoting where the motion trackers were in relation to their ascribed landmarks.

Abnormal Scapular Motion

Kibler and colleagues (2002) described three types of abnormal scapular motion. Type I is inferior angle dyskinesia. In this, the inferior medial scapular border tilts dorsally becoming prominent during arm motion. Medial border dyskinesia (type II) reflects dorsal motion of the entire medial scapular border. Type III is superior border dyskinesia in which shoulder shrugging initiates the motion. All three of these reflect a loss of smooth control of the scapula and each takes place around one of the three axes of scapular motion (p. 551).

Several authors have suggested causes for the loss of muscular control of the scapular stabilizers. Ludewig and Cook (2000) noted that decreases in serratus anterior activity, increases in upper trapezius activity, or any imbalance between the upper and lower trapezii have been linked to kinematic changes (p. 278). Kibler (1998) observed that injury to the long thoracic nerve or spinal accessory nerve can result in alterations and inhibitions of the serratus anterior or lower trapezius. Furthermore, direct trauma to either of these muscles can result in inhibition of normal function (p. 327).

Abnormal positioning of the scapula and inappropriate motion, reflected by the tilts described in the three types of dyskinesia, predispose the individual to further injury. Lukasiewicz et al. (1999) described the alterations in scapular positioning in individuals with impingement syndrome. Schmitt and Snyder-Mackler (1999) further described this relationship observing concomitant serratus anterior weakness in a case of primary subacromial impingement in addition to middle and lower trapezius weakness. In this case the patient had type I inferior medial border winging. This is explained by the loss of

control of the serratus and lower trapezius which function to stabilize the inferior angle of the scapula against the thorax at rest and as the scapula upwardly and downwardly rotates (p. 32).

Kibler (1998) elaborated on the spectrum of pathologies that can result from the dyskinetic scapula. First, lack of full protraction increases the deceleration forces during throwing which can result in both micro- and macrotrauma to the external rotators. Furthermore, a loss of protraction results in a loss of the bony continuity of the glenohumeral joint as the anterior humeral head extrudes further anteriorly during arm deceleration. In this position, there is an anterior opening of the glenohumeral joint and the static structures of stabilization, the glenohumeral ligaments, incur significant stress leading to laxity and a predisposition toward glenohumeral instability (p. 329).

Excessive protraction reduces the scapula's ability to elevate and upward rotate to allow for acromial elevation and thus increases the possibility of impingement syndrome. Impingement can result in rotator cuff tendinitis, glenohumeral instability, and a perpetual cycle of exacerbation as the lower trapezius and serratus anterior are forced to compensate for, and are ultimately inhibited by, the loss of range of motion and strength decrements resulting from inefficient supraspinatus function (Kibler, 1998, p. 329).

Significantly, scapular dyskinesis leads to a breakdown of the kinetic chain. Force generated in the lower extremity and torso can no longer be appropriately transmitted to the shoulder, arm, and hand as the scapula is no longer a stable base of muscle attachment and force transmission. The ultimate result of this loss of controlled motion is greater muscle activation and force generation being required of the intrinsic and extrinsic

shoulder muscles as well as the muscles of the arm and hand in order to compensate for the lost kinetic force. These muscles thus become prone to fatigue, overuse, strains, and rupture (Kibler, 1998, p. 329).

Evaluation of the Dyskinetic Shoulder

Evaluating the quality of scapular motion during the clinical examination of a shoulder is of paramount importance. It is also incredibly difficult because the bone and its articulation against the thorax are obscured by the overlying musculature. Body composition can also potentially interfere with viewing the scapula, particularly in individuals with larger BMI values. Several authors have suggested methods for observing and classifying abnormal scapular motion.

Traditional evaluation of scapulothoracic function involves visual inspection and manual muscle testing. Warner and colleagues (1992) noted that these measures may be insufficient to detect the gravity of weakness in the musculature. Additionally, qualitative visual inspection may be difficult on individuals with well-developed scapular musculature or excessive body fat which limits visualization of the scapula.

Warner et al. (1992) proposed a three-dimensional model using Moiré topography to detect these subtle dysfunctions. In this method, the participant “is positioned behind a grid of horizontal beams of light created by a point light source. The line shadows cast by the grid conform to the surface topography of the subject” (p. 192). For their experiment, participants being treated for impingement or glenohumeral instability were compared to control participants with no history of shoulder dysfunction. Both static positions and dynamic shoulder motions were evaluated and during the dynamic assessment on the last

of 10 elevation/lowering reps, a photograph was taken during the final 60° to 30° of motion.

Both the impingement and instability groups demonstrated asymmetries, increased topography, and winging during the test. Though the control group demonstrated some asymmetry, the differences between the impingement and instability groups versus control were significant. The test showed the Moiré topography method to be sensitive to subtle changes in the kinematics of the scapulothoracic motion, more so than is available in traditional testing models (Warner et al., 1992).

However, the use of Moiré topography and this particular model pose challenges in the clinical setting. First, the apparatus required to generate the topographic shadows is not readily available to most clinics nor is a room specifically designated for this purpose, as a dark room is necessary. Secondly, to be able to distinguish normal from abnormal in the topographic map requires extensive training and subtle differences are not necessarily apparent during visual observation of the subject, particularly during dynamic motion. Often, photographic evaluation of the individual is required to detect topographic alterations.

Photographic evidence during the Moiré exam also poses problems. The authors note that the particular range chosen for photography reflects the point at which the lower trapezius and serratus anterior are maximally working eccentrically to stabilize the scapula against the thorax and thus deficiencies are most susceptible to demonstrate insufficiencies. However, without a camera capable of rapid shutter speed, it is likely that the alterations would be missed by the camera. Furthermore, this range neglects the

possibility of type III dyskinesia, which is demonstrated by an initial shoulder shrug during arm elevation. Despite the sensitivity of Moiré topography to contour abnormalities during scapular motion, the drawbacks make it impractical and unfeasible for clinically assessing scapular dyskinesia.

The lateral scapular slide test (LSST) is a semidynamic test described by Kibler (1998). The positions of the scapulae are measured relative to a fixed position on the spine. The arm is moved into three positions for measurement: relaxed at the sides, hands on the hips with the fingers anterior and thumb posterior and maximal internal rotation at or just below 90 degrees of elevation (Kibler, 1998). Kibler asserted that these positions load the posterior scapular stabilizers and that weaknesses of these muscles will be reflected in asymmetric positioning. A side-to-side difference greater than 1.5 cm in any of the three positions results in a positive dyskinesia diagnosis.

Odom and colleagues (2001) have criticized the utility of the LSST. They found mid to high intertester reliability (range: 0.43 to 0.79) and high intratester reliability (range: 0.52 to 0.8). The population of participants consisted of subjects with and without additional shoulder impairments. Given the association between scapular dyskinesia and other shoulder injury, Odom et al. calculated sensitivity and specificity for the three test positions and two thresholds, 1 cm and 1.5 cm bilateral difference, using a priori physician diagnosis of injury as the reference criterion. The authors found low sensitivity (range: 35% - 43%) and low specificity (range: 48% - 56%) for the LSST, suggesting high false positive and negative rates. The authors conclude that “[s]ensitivity and specificity of the LSST are unacceptably low, with the LSST performing little better at

classification than chance alone (Odom et al., 2001).

Kibler et al. (2002) proposed an entirely clinical method for assessing scapular motion. The authors determined the three classifications of abnormal scapular motion described earlier. A fourth classification represented symmetric scapular motion and thus normal motion. These patterns of motion were verbally and visually presented to two physical therapists and two physicians experienced in orthopedics. Participants in the study were videotaped for later review and were not assessed at the time of testing. In the testing procedure, participants performed three repetitions of bilateral arm elevations and lowering in both the scapular plane and abduction at a rate of 45°/s.

The clinicians were then asked to evaluate each participant as to which of the four scapular patterns was predominant. One of the physicians and one of the physical therapists then reviewed the same videotape at a later date to establish intratester reliability. The agreement between the member as well as the intratester reliability was moderate, below 0.50 with the exception of the intratester reliability for the physician at 0.59. However all of the reliabilities were determined to be statistically significant (Kibler et al., 2002).

This method, too, imposes certain challenges to use and the methodology of the testers contributes to these challenges. This experiment aimed to establish the reliability of experienced clinicians to classify an individual into one of four categories based on the predominant pattern displayed. However, as the authors note, “patients may show combinations of the patterns because of the scapular movement in 3 dimensions.” The authors provide no data on the reliability of classifying participants into the normal

category versus any abnormal category. It may be that the clinicians were adept at identifying individuals with abnormal motion but disagreed as to the type. It may be that the clinicians were not adept at identifying the individuals with abnormal motion at all. The data provided do not elucidate this.

Additionally only six individuals with normal shoulder range of motion and no history of shoulder injury were used. The small “control” group may have affected the outcome. Agreement may have been higher if more “normal” participants had been viewed.

The procedure itself also seemed counterproductive to demonstrating abnormal scapular kinematics. The authors limited the repetitions to three each of the abduction and scaption and make specific note that this was done “in a counterbalanced order to prevent fatigue.” The repetitions were also performed without external weighting. As demonstrated by McQuade and Smidt (1998) and McQuade et al. (1998), fatigue or external loading are sufficient to induce alterations in the scapulohumeral rhythm. The clinicians may have been better able to determine predominant scapular patterns if a fatiguing or externally loaded protocol was utilized. The moderate reliability coupled with the methodological concerns related to this study make the method for evaluating scapular dyskinesis in Kibler et al. (2002) tenuous at best.

McClure and colleagues (2009) have similarly developed a visual model for classifying scapular dyskinesis which improves upon the Kibler et al. (2002) method's shortcomings. For McClure et al. (2009) a standardized training program was developed which included operational definitions, photographs of normal and abnormal scapular

motion as well as video which represented the various possible abnormalities.

Participants were overhead, Division 1 collegiate athletes with no specified control condition. In the testing procedure, participants performed five repetitions each of bilateral shoulder flexion and bilateral shoulder abduction with dumbbells matched to body weight categories. The participants were videotaped for later review and assessed during the testing procedure. Examiners rated both flexion and abduction independently and for both shoulders independently. Ratings were normal, subtle or obvious per motion and then either normal, subtle, or obvious per shoulder depending on the ratings given to the individual motions (McClure et al., 2009).

Interrater reliability for the McClure et al. (2009) system was higher than for the system outlined in Kibler et al. (2002) ($\kappa = 0.54-0.57 > \kappa = 0.31-0.42$) and was consistent for live versus videotape raters. This method has several advantages over the Kibler et al. method. First, examiners evaluate the scapulae independently of each other which reduces the reliance on asymmetry as a criterion which may be present at rest and persistent through motion and thus would produce inconsistent results. Second, weights and more repetitions are utilized which are more likely to induce fatigue and exacerbate symptoms of scapular dyskinesis. This makes it easier for the examiner to clearly see evidence of abnormal motion.

McClure et al. (2009) also did away with the necessity to grade the type of dyskinetic motion. Subtle or obvious ratings simply required the presence of one or more of the types of dyskinesis without specifying the type. As noted in the shortcomings of Kibler et al. (2002), greater dissent regarding scapular dyskinesis is present when raters

are required to be exact about the type of abnormal motion rather than assessing whether or not there is any abnormal motion.

Tate and colleagues (2009) further evaluated this method by comparing those individuals rated as either normal or obvious scapular dyskinesis with three-dimensional kinematic data. The three-dimensional tests replicated the original test procedure but only consisted of three repetitions each. Sensors were placed on the manubrium, humerus, and scapula. Details of scapular placement only state that “the receiver was applied to a custom-made, adjustable scapular tracking jig” (p. 167). No pictures or further descriptions of the jig were given so it is unclear how well it mimicked scapular motion. However, the authors reported 0.15° accuracy of the entire device to have been previously verified.

The authors found that the dyskinesis-rated group demonstrated less upward rotation, less clavicular elevation, greater posterior tilting, and greater protraction than the normal-rated group at rest and during motion. They also noted increased resting internal rotation with subsequent greater relative external rotation during motion in the dyskinesis group versus the normal group. The authors also determined that pain as demonstrated on the Penn Shoulder Scale was not predictive of the presence of scapular dyskinesis (Tate et al., 2009).

The three-dimensional kinematic data from Tate et al. (2009) represents a quantifiable confirmation of the findings of McClure et al. (2009). Individuals rated as having an obvious abnormality demonstrated via electromagnetic tracking an actual difference in scapular kinematics. This confirmation is significant because it reflects

objective confirmation of a qualitative system for judging scapular dyskinesis. The test devised in McClure et al. has sufficient reliability to be utilized clinically and reflects an actual deviant process.

Uhl et al. (2009) went on to evaluate the diagnostic accuracy of this dichotomous method when compared against the four part method from Kibler et al. (2002). Uhl et al. (2009) reported results in terms of agreement, statistical reliability, sensitivity, specificity, positive and negative likelihood ratios, and accuracy. Clinical observation was compared against three-dimensional electromagnetic modeling for sensitivity, specificity, positive and negative likelihood ratios, and accuracy. Subjects for the study were both symptomatic patients of one of the authors as well as asymptomatic participants from the community.

Uhl et al. (2009) found comparable agreement for the 4-type clinical assessment and the dichotomous method, 61% and 79%, respectively. They also found very similar reliabilities, κ correlations of .44 and .41, respectively. Overall, the dichotomous method demonstrated greater sensitivity (74%-78%), greater positive predictive value (76%-78%), comparable negative predictive value (27%-40%), and comparable accuracy (64%-66%) to the previously described 4-type method (10%-47%; 20%-58%; 50%-78%; and 45%-64%, respectively) while specificity was decreased (31%-38% versus 62%-94%). The authors state that “[t]his indicates that the yes/no assessment method decreases the risk of false-negative findings by better identifying subjects who truly have scapular dyskinesis” (p. 1246). Furthermore, “[t]he specificity of the yes/no method was 30%, indicating that there is a higher risk of false-positive findings. This combination of values

indicates that the yes/no method is a good screening tool in the shoulder evaluation process and provides greater agreement among clinicians (inter-rater reliability) in their observational assessment of scapular dyskinesis” (p. 1246).

Interestingly, the authors also found a high incidence of asymmetric motion in both the symptomatic and asymptomatic groups, 71%-77% for all subjects. According to the authors, “[t]his indicates that the presence of asymmetry should not be the sole criterion determining the significance of scapular dyskinesis” (p. 1246). While this is certainly true and bilateral differences may only be useful in the presence of other clinical symptoms, the high occurrence has definite uses in further research into scapular dyskinesis. Whereas previous studies have relied on symptomatic populations, particularly athletes, recruitment pools can be larger when assessment methods specifically targeting identifying dyskinesis are being researched. Researchers will be able to draw on athletes and non-athletes, symptomatic and asymptomatic individuals because the presence of dyskinesis may manifest in any of these populations.

An in-publication doctoral dissertation by Priscilla Dwelly (2012) examined the ability of athletic training students to evaluate scapular dyskinesis. In this study, 41 graduating senior athletic training students were asked to evaluate videos of 15 individuals performing actions similar to those found in McClure et al. (2009) and Uhl et al. (2009). The students were asked to evaluate these participants using the 4-part dyskinesis rating system from Kibler et al. (2002). The ratings provided by the students were compared against the ratings of six experts, those who had “demonstrated with peer reviewed publications an understanding and competence in the cause, evaluation, or

rehabilitation of scapular dyskinesis” (Dwelly, 2012). Using the expert rating as “correct,” Dwelly found that the students answered correctly 81% of the time. Reliability was calculated at $\kappa=0.32$, consistent with the reliability found by Kibler (2002).

At 4-week follow-up, a second questionnaire was sent to the 41 athletic training students. It contained seven out of the 15 videos to test intra-rater reliability. Nineteen responded to the questionnaire and intra-rater reliability was calculated at $\kappa=0.45$, demonstrating moderate test-retest reliability. These findings concur, as the author noted, with recent studies on visual methods for evaluating dyskinesis (McClure et al., 2009; Uhl et al., 2009).

Diagnostic Accuracy

When selecting clinical tests to utilize, the athletic trainer should be mindful of the degree to which the results of the test reflect the reality of the presence or absence of a condition. There are a number of values which can be calculated to determine the

Table 1

2x2 Contingency Table for Calculating Diagnostic Accuracy

Clinician Rating	Physician Rating		
	Dyskinetic	Normal	
Dyskinetic	True positive (a)	False positive (b)	PPV= $a/(a+b)$
Normal	False negative (c)	True negative (d)	NPV= $d/(c+d)$
	Sensitivity = $a/(a+c)$	Specificity = $d/(b+d)$	Accuracy = $a+d/(a+b+c+d)$

Note. PPV is positive predictive value. NPV is negative predictive value.

diagnostic accuracy of a test. To calculate these measures, a 2x2 contingency table is constructed. Along the top, a reference standard, or gold standard, is used. A positive reference standard forms the left column and a negative reference standard forms the right column. Rows are created by the clinical test's results. Positive tests are the top row while negative tests are the bottom row. Thus the 2x2 contains the results reflected as true positives, false positives, true negatives and false negatives. An example is provided in Table 1.

To gain an overview into the utility of a test, the overall accuracy can be calculated. Overall accuracy is a percentage and reflects the total number of “true” tests contained in the contingency table compared to the total number of tests conducted. This is a simple percentage though and as Cleland (2005) notes, “[t]he accuracy of a diagnostic test should not be used to ascertain the diagnostic accuracy of that test because overall accuracy can be a bit misleading. The accuracy of a test can be significantly influenced by the prevalence ... in the population at a given time” (Cleland, 2005). Thus while overall accuracy can give a rough overview, it should not be the sole statistic used.

Looking at the 2x2 table vertically gives us results in reference to the gold standard or reference criterion's assessment. Sensitivity is measured using the positive reference criterion column. It is the percentage of true positive tests out of all results when the condition is present, that being true positives and false negatives. This reflects the value of a negative clinical test at ruling out the condition. Though it is based on the true positive rate, sensitivity does not influence the interpretation of a positive result, as it may be contaminated by the false positive rate (Fritz & Wainner, 2001). Rather it reflects

the importance of a negative result at ruling out the presence of the condition. High sensitivity means that a negative result is most likely to be true.

Specificity is calculated from the right-hand column and is a percentage of true negatives out of all of the results in light of a negative by gold standard. This is the opposite of sensitivity. Specificity is a measure of the value of a positive clinical test result. When specificity is high, the false positive rate is low and thus a positive clinical test result likely reflects the condition being present. Sensitivity and specificity values range from 0, no true positives or negatives respectively, to 1. Fritz and Wainner (2001) note that few tests have both high sensitivity and specificity. It therefore becomes incumbent on the clinician to decide whether to trade off high sensitivity in deference for high specificity, thereby accepting more false positives, or vice versa.

Looking horizontally at the contingency table defines the interpretation of clinical tests in light of the reality of the condition. The top row, a positive clinical test, allows for the calculation of the positive predictive value (PPV). The bottom row, a negative clinical test, enables the determination of the negative predictive value (NPV). These two values refer to the percentage of true positive or negative results out of all positive or negative results, respectively. High values for these can be powerful indicators that future positive or negative clinical results indicate the true presence or absence of the condition. However, this must be tempered with knowledge that the prevalence of the condition can significantly alter these values (Fritz & Wainner, 2001).

Finally, utilizing sensitivity and specificity values, likelihood ratios can be calculated. These ratios determine the posttest shift in probability that the clinical

assessment reflects the disorder. A positive likelihood ratio greater than 1 suggests a shift in favor of the presence of the condition. A negative likelihood ratio less than 1 suggests a shift in favor of the absence of the condition. Combined with pretest probability, the values can be used to shift the probability of that the result of a clinical examine reflects reality. In general it is desirable for positive LRs to be greater than five and negative LRs to be close to zero in order to demonstrate large shifts. Values centering around 1 indicate minimal shift and thus continued uncertainty (Fritz & Wainner, 2001).

Research Considerations

There are several discernible shortcomings in the research on the analysis of scapular dyskinesis in the athletic training domain. The first, and most glaring, is simply the lack of research utilizing athletic trainers as clinicians. In McClure et al. (2009), two athletic trainers viewed videos of participants. In Uhl et al. (2009), one of the evaluators was an athletic trainer. And most recently, Dwelly (2012) utilized 41 athletic training students. However, as Dwelly notes, the evaluation of scapular dyskinesis is a necessary component of the evaluative process and a part of the competencies defined by the National Athletic Trainers' Association's competencies. It seem appropriate, therefore, that more thorough investigations into the use of scapular dyskinesis evaluative methods in the hands of athletic trainers be undertaken.

Second, many of the studies published utilize video-taped participants. While video-taping allows for mass distribution and requires little coordination of subjects and clinicians in order to conduct testing, it is possible that something is lost in the translation from live to video.

Finally, throughout the research on scapular dyskinesis is the emphasis on the overhand athlete as the only potential client in whom the clinician is likely to see such alterations. As Burkhart et al. (2003) observed, abnormal scapular motions which the authors term the SICK scapula are frequent in throwers in particular because of the high forces generated in the musculature of the shoulder girdle. These forces have two specific results. The first is microdamage to the posterior musculature and posterior glenohumeral capsule resulting in a loss of internal rotation. This accompanies a naturally “dropped” shoulder and increased resting protraction. These three symptoms alone are sufficient to begin the cascade of problems described earlier resulting in rotator cuff injuries, glenohumeral laxities, and the potential for labral insult.

However, throwers and overhead athletes are not the only individuals who can present with scapular abnormalities. The computer age has resulted in a resting posture for the general populace in which the scapulae are already protracted. Without sufficient stimulus to strengthen the serratus anterior and lower trapezius, muscles already prone to weakness, the population at large is at risk for scapular dyskinesis and potential shoulder injuries without engaging in high energy, overhand athletics.

CHAPTER III

METHODS

To address the first research purpose of whether graduate assistant athletic trainer can utilize the yes/no SDT (McClure et al., 2009) as reliably as experts, ratings of dyskinesia utilizing this system were compared to previously published reliability data using the system. To address the second research question of whether the yes/no SDT in the hands of a graduate assistant athletic trainer is a useful tool in upper extremity assessments, sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, and accuracy were calculated against a physician's assessment of the same participant pool. Prior to data collection this project was reviewed and approved by the Institutional Review Board (IRB) of Utah State University. All testing was conducted in the Dale Mildenerger Athletic Training Room at Utah State University.

Participants

Participant Recruitment

Participants in this study were recruited in two phases. First, requests for participation were made to three undergraduate classes in the Health, Physical Education, and Recreation department at Utah State University. Participants ($n=7$) were included in the study if they could reach a minimum of 120 degrees of forward flexion and scaption; did not have bilateral shoulder pain at the time of testing; had no history of fracture of the scapula, humerus, or clavicle; and had no history of injury to the long thoracic nerve,

spinal accessory nerve, or any of the cervical nerve roots. Additionally, the clinicians used for assessing dyskinesia were asked to participate in the study. Of the clinicians ($n=10$), nine volunteered to be subjects for evaluation. They were accepted as subjects following the same inclusion criteria. Finally, the author also engaged in the study procedures to serve as a participant. In all, 17 participants underwent examination for scapular dyskinesia.

Participant Characteristics

The participants ($n=17$) in this study were all college students. The mean age of the participants was 23.9 years with a range of 21-28. None of the participants was actively engaged in Division 1 athletics. The sports previously played are detailed in Figure 2. One participant had not engaged in any high school or recreational collegiate athletics.

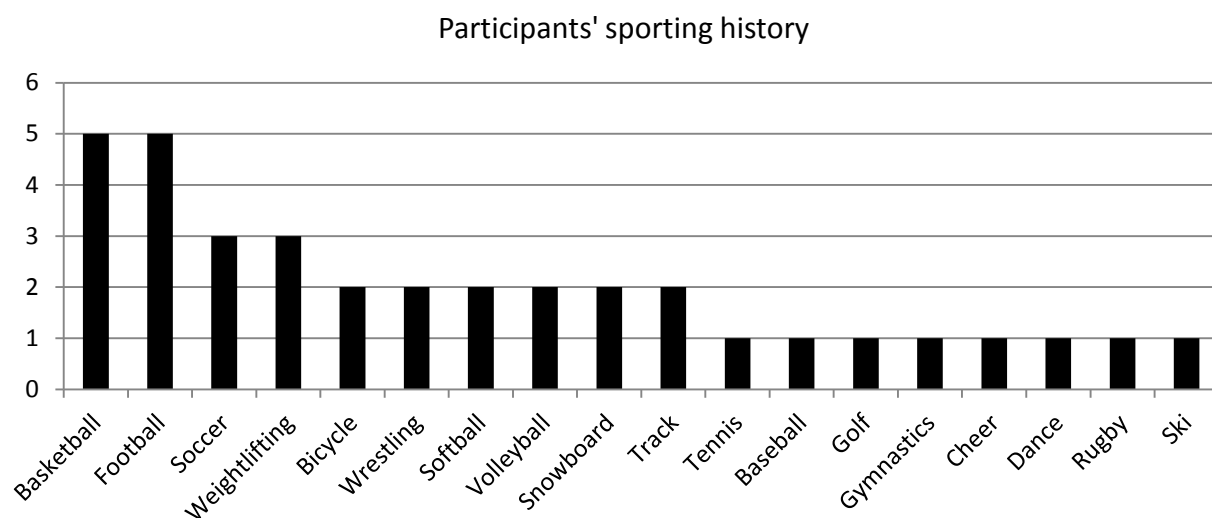


Figure 2. Number of participants engaging in select sports during high school or college.

Clinicians

The clinicians of interest were all graduate assistant athletic trainers ($n=10$) at Utah State University. All of the clinicians had been certified by the Board of Certification prior to this study. The average length of certification was 1.3 years (range = .75 years to 1.5 years). Two pairs of the clinicians came from the same undergraduate programs. The remaining six clinicians came from other programs, giving a diverse set of prior educational and clinical experiences.

Procedures

Training

Each of the clinicians underwent a training protocol for the education and identification of scapular dyskinesis consistent with definitions used in the initial SDT study conducted by McClure et al. (2009). This training module consisted of definitions, images, and video examples of both normal and dyskinetic motion presented via portable document format (PDF). The training module was developed by Dr. Philip McClure and is available on his faculty page at Arcadia University (<http://www.arcadia.edu/academic/default.aspx?id=15080>). This is the same training received by the clinicians in McClure et al. study.

Each of the clinicians in the present study viewed the PDF on either a laptop or desktop computer. At the end of the training module there was a built-in post-test. If the clinician did not score 100% on the post-test he or she was encouraged to review the material in the training until 100% accuracy was attained. They were instructed not to discuss the training with one another. Raters were also instructed not to discuss the results

of any participant during or after the testing periods until all data had been collected to minimize the influence of group assessment on individual ratings.

SDT Test

Prior to arriving for testing, all participants were informed of the study procedures and exclusionary criteria. Men interested in completing the testing were asked if they would be comfortable removing their shirts for the duration of the testing to allow for visualization of the scapulae throughout the full range of motion. Women interested in participating were asked to wear a spaghetti strapped shirt, jogging or sports bra, or other article of clothing which would allow for full visualization of the scapulae throughout the full range of motion.

All of the participants signed the IRB approved Informed Consent form prior to participation (see Appendix A). Participants were asked if they had a history of clavicular, humeral, or scapular fracture or long thoracic, spinal accessory, or cervical nerve injury or impairment. All participants who had previously agreed to be included in the study responded negatively to these questions.

Each participant was then evaluated by Dr. Trek Lyons, the team physician for Utah State University athletics. His evaluative procedure is listed in Appendix B. For the physician, each participant was rated as either normal or dyskinetic and a determination was made about the side in the event of dyskinesia.

Following the physician's assessment of scapular motion, the participant was instructed in the procedure for the SDT. The participant was instructed to hold a pair of dumbbells corresponding to body weight (3 lbs for those weighing less than 150 lbs; 5 lbs

for those weighing greater than 150 lbs). With the scapulae exposed, the participant performed five repetitions of forward shoulder flexion to 120 degrees at a pace of 3-seconds ascending, brief pause, and 3-seconds descending. This timing was then repeated for five repetitions of scaption. During this time, five clinicians viewed the motion of the scapulae. Each rater then marked a score sheet (Appendix C) indicating whether the participant had either normal or abnormal scapular motion. If dyskinetic motion was detected, the clinician then indicated whether it was observed on the left side, right side, or both sides. The participant was then given a 2-3 minute break and the testing procedure was repeated for the other five clinicians. Following testing procedures, demographic information including age and past competitive sport involvement was obtained. Afterwards, participants were dismissed. The total time commitment for participation was approximately 20 minutes.

Statistical Analysis

Reliability

To determine the inter-rater reliability of the SDT test, Gwet's AC1 coefficient was calculated (Gwet, 2008). Typical methods for calculating reliability rely on the kappa statistic. Cohen's kappa is designed for two raters. Fleiss' kappa and several other statistical derivatives have been developed for use with multiple raters. However, there is a noted paradox when using kappa statistics: When percent agreement is high, the kappa statistic is artificially suppressed due to the homogeneity of responses (Warrens, 2010). Accordingly Gwet developed the AC1 statistic as a more robust coefficient in the presence of high rater agreement (Gwet, 2008). Given the high level of agreement shown

in Uhl et al. (2009) and Dwelly (2012), the AC1 statistic is appropriate.

To calculate Gwet's AC1, AgreeStat2011.1 was used. AgreeStat2011.1 is a Visual Basic Application (VBA) script designed to run inside Microsoft Excel. It provides numerous multiple rater reliability coefficients: Conger's kappa, Gwet's AC1, Fleiss' kappa, Krippendorff's alpha, Brennan-Prediger, and percent agreement. In addition, it gives the standard error and 95% confidence intervals per coefficient. It also allows for missing data points. The ability to calculate missing data points was essential because 9 of the 10 raters served as subjects and were therefore unable to assess on one trial.

Data were analyzed under two conditions. The first was whether or not the rater rated the participant as normal or dyskinetic. The second was the agreement about which side was identified as being dyskinetic. Previous studies (McClure et al., 2009; Uhl et al., 2009) demonstrated kappa values ranging from $\kappa = 0.41$ -0.57. Accordingly, a kappa score exceeding $\kappa = 0.4$ is expected.

Diagnostic Accuracy

Diagnostic accuracy for the SDT test was assessed by calculating sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy. These were calculated in Microsoft Excel using the 2x2 contingency tables described in Cleland (2005). Refer to Table 1 for the contingency table.

In each case, physician assessment was considered to be the reference standard. True positives and negatives were those in which the physician and clinician agreed about the assessment of the participant. False positives were those in which the physician rated the participant to have normal motion while the clinician rated the participant to have

dyskinetic motion. False negatives were those in which the physician rated the participant to have dyskinetic motion while the clinician rated the participant to have normal motion.

These indices of diagnostic accuracy were calculated under the following conditions: The physician rated the left side to be dyskinetic, the physician rated the right side to be dyskinetic, and the physician rated both sides to be dyskinetic. To get an overall examination of the indices, the data were consolidated into normal or abnormal, regardless of side and the contingency table was constructed.

CHAPTER IV

RESULTS

The physician evaluated 16 of the 17 participants to have dyskinetic motion in one or both scapulae. Of the 16, four were evaluated with left sided only dyskinesia. Five were evaluated with right-sided only dyskinesia. Seven were evaluated to have dyskinesia in both shoulders. Only one was evaluated to have bilaterally normal motion.

To evaluate the reliability of graduate assistant athletic trainers' use of the yes/no SDT, Gwet's AC1 statistic and percent agreement were calculated on two sets of data. The first set consisted of the ratings by the athletic trainers as to whether motion, when taken as a whole, was normal or dyskinetic, irrespective of side (Table D1). The second set consisted of the ratings when sided-ness was taken into account. In four instances, raters observed dyskinesia to be present but could not determine which side was normal and which was dyskinetic or if both were dyskinetic. This data set is presented in Table D2.

For the first data set, the clinicians agreed 72% of the time. The raters displayed moderate reliability ($AC1 = 0.48$, $p < 0.0025$, 95% CI [0.147, 0.812]). When sided-ness was taken into consideration, the clinicians agreed 54% of the time and displayed moderate reliability ($AC1 = 0.43$, $p < 0.0001$, 95% CI [0.242, 0.632]).

To determine the diagnostic accuracy of the test, sensitivity, specificity, PPV, NPV, positive and negative likelihood ratios, and accuracy were calculated on a present or absent basis and a side-by-side basis. Table 4 shows the ratings by the physician and clinicians of only those participants for whom left-sided dyskinesia was observed. Tables

D3 and D4 give this information for right-sided only and both sided dyskinesia, respectively. Table D5 shows participants in whom the physician or clinician detected dyskinesia, regardless of side. Table 2 summarizes the diagnostic accuracy by condition in terms of true and false positives and negatives. Table 3 gives the tests for diagnostic accuracy per condition as well as a composite of the diagnostic accuracy across conditions.

Overall, when the clinicians were asked only to identify whether dyskinesia was present or not, regardless of side, there was moderate sensitivity (66.2%), moderate specificity (50%), high positive predictive value (95.2%), low negative predictive value (8.9%), and low positive and negative likelihood ratios (1.325 and 0.675, respectively) with an accuracy of 65.2%. When asked to choose which side demonstrated dyskinesia, clinicians had low sensitivity (34.4%), high specificity (85.2%), moderate PPV (51.5%), moderate NPV (74.1%), and moderate accuracy (69.4%). The PLR (2.333) and NLR (0.769) are both considered to be low.

Table 2

Diagnostic Accuracy by Count

Condition	True		False	
	Positive	Negative	Positive	Negative
Physician and clinician agree				
Left only	10	113	9	29
Right only	13	102	11	35
Both only	29	68	29	35
Total	52	283	49	99
Physician and clinician detect dyskinesia				
Any side	100	5	5	51

Note. The first three conditions indicate times when the physician rating and the clinician rating matched exactly. The fourth condition indicates times when both the physician and clinician observed dyskinesia, regardless of side. True and false reflect physician rating as a reference.

Table 3

Diagnostic Accuracy

Condition	Sensitivity	Specificity	PPV	NPV	Accuracy	PLR	NLR
Physician and clinician agree							
Left only	0.256	0.926	0.526	0.796	0.764	3.476	0.803
Right only	0.271	0.903	0.542	0.745	0.714	2.782	0.808
Both sides	0.453	0.701	0.500	0.660	0.602	1.516	0.780
Exact Matches	0.344	0.852	0.515	0.740	0.694	2.333	0.769
Physician and clinician detect dyskinesia							
Any dysk.	0.662	0.500	0.952	0.089	0.652	1.235	0.675

Note. PPV is positive predictive value. NPV is negative predictive value. PLR and NLR are positive and negative likelihood ratios and are expressed as ratios.

CHAPTER V

DISCUSSION

Evaluation of scapular dyskinesis is a key component of any upper extremity evaluation. Athletic trainers are expected to be able to perform such screens and to date have had few adequate tools for scapular evaluation. Since 2009, the scapular dyskinesis test proposed by McClure et al. (2009) has come to light as a potential front-runner in the upper extremity screen. This test has shown moderate inter- and intra-tester reliability and agreement which is, if not exceeding then on par with, other clinical tests for scapular dyskinesis. Furthermore, this test has been validated by Tate et al. (2009) and Uhl et al. (2009) against three-dimensional kinematic data to demonstrate that clinicians can detect scapular motion irregularities.

This study confirms much of what has previously been found and published. The certified athletic trainers in this clinician pool, after only utilizing the training module from past studies, were in agreement that there was some degree of dyskinetic motion 72% of the time. This is highly consistent with the findings of Uhl et al. (2009), whose clinicians were two experts in the field of scapular dyskinesis, where the agreement was 79%. Furthermore, the reliability found in the present study was in the moderate range at 0.48 but on par with the kappa statistic from Uhl et al. at $\kappa = 0.41$ and McClure et al. (2009) at $\kappa = 0.55$.

There is a decrement in the agreement among the clinicians in the present study when asked to determine which side exhibited dyskinetic motion. They still agreed on over half (54%) of the assessments without a significant decrease in reliability (0.43).

Even with the decrease, it is reasonable to believe based on this, and previously published works, that when using a system that relies only on whether there is normal or abnormal motion, not what kind, clinicians of varying degrees of experience can be taught to use this system and use it as reliably as experts.

The diagnostic accuracy measures are less convincing. Sensitivity and specificity indicates how correctly clinicians can rule in or rule out individuals with scapular dyskinesis, in this case. Sensitivity indicates the strength of a normal evaluation not being a false negative. In this study, it is much higher when the clinicians are only asked to identify whether dyskinesis is present or not rather than identifying a particular side (66.2% versus 34.4%). This indicates that clinician ratings of normal are likely to be correct approximately two-thirds of the time if the clinician is only interested in detecting whether or not there is some measure of abnormal motion.

Specificity indicates how confidently the clinician can state that a positive rating truly is a positive, rather than a false positive. Here, the specificity is enhanced when the clinician is asked to determine which side is dyskinetic rather than just whether dyskinesis is present at all (85.2% versus 50%). As specificity approaches 1.0, the clinician's determination of a positive test gains strength because of the low false positive rate. Thus the determination of side decreases the false positive possibility by 35%.

Positive and negative predictive values display opposite relationships in this particular study as well. PPV, a measure of positive test being a true positive, is significantly greater when the clinician determines whether or not dyskinesis is present rather than choosing a side. At 95% PPV, the false positive rate is limited to only 5%. The

opposite is true in this case for negative predictive value. The NPV is a measure of a negative test truly being negative. NPV is more diagnostic when the clinician chooses a side rather than simply identifying when dyskinesia is present. There is a significantly higher true positive to false positive rate when side is ignored and a higher true negative to false negative rate when side is determined.

The positive and negative likelihood ratios are post-hoc measures which indicate the strength of the probability that a positive or negative clinical result reflects reality after that result has already been obtained. Because the PLR and NLR hover near 1, and certainly less than 5, under both conditions, they are of little value to the clinician. The accuracy regardless of whether side is considered or not is approximately the same, between 65% and 70%.

The burden on the clinician in using these data is to decide what is more important: false positives or false negatives. With a false positive in a symptomatic individual, the clinician really cannot do much harm in devising a rehabilitation program which maintains the strength and firing pattern of the scapular muscles. Unless a muscle, or two muscles, is totally neglected, there is really no foreseeable risk. However, false negatives pose a much more serious problem. The clinician risks perpetuating the injury which initiated the assessment by not correctly evaluating a scapular motion error. Instead of rehabilitating fully, the patient may continue to be at risk for future shoulder complex injury.

These diagnostic accuracy measures suggest a combination of the two clinical options to strengthen outcomes. To avoid false negative results, it is more diagnostic to

assess whether dyskinesia is present at all rather than determining side. By doing so, the sensitivity and positive predictive values are enhanced. These values compare well with Uhl et al. (2009) with sensitivity being slightly lower in the present study but PPV being significantly higher as well as specificity being higher. Accuracy between the two studies is comparable. The current results also suggest that in light of a decision of dyskinesia, regardless of side, it is of merit to then determine a side. In doing so, the specificity is further enhanced and side-by-side ratings of dyskinesia can likely be stated to be truly positive.

Limitations

There were a number of limitations of the present study. The first was participant recruitment. Due to the fact that all of the clinicians were employed with the athletics teams at Utah State University, athletes were not specifically recruited. The reasoning was to reduce the risk of bias due to previous evaluation or knowledge of medical history. Prior studies' participant pools included at least a subset of overhand athletes. For this study, the lack of athletes in the recruitment process may have compelled such a small sample size.

Given the high prevalence of dyskinesia in asymptomatic populations reported by Uhl et al. (2009), college classes were used as recruitment sites. Though verbal interest was high, actual commitment to come in for testing was meager ($n=7$). Given the overall small sample size ($n=17$) and diverse levels of physical activity among the participants, the overwhelming prevalence of dyskinesia as determined by physician assessment may have played a significant role in decreasing the overall accuracy of the SDT, as suggested

by Alberg, Park, Hager, Brock, and Diener-West (2004). Alberg et al. state that “overall accuracy is most problematic as a measure of test validity when the prevalence is very low or very high.... When prevalence is high, overall accuracy more closely resembles sensitivity” (p. 462).

The meager response from recruitment can likely be attributed to the restrictiveness of the testing schedule. Though the testing procedure itself took no more than approximately 20 minutes, the time and day were constrained to Wednesday evenings after 7. This was due to schedule conflicts in trying to coordinate 10 graduate assistants and a physician to all be available at the same time for in-person testing. This time frame may have also confounded the data somewhat as it was after the end of the workday and the clinicians may have been ready to be done and not as focused as possible.

Due to a small response by the potential participant pool as well as schedule conflicts, testing took place on two evenings which were separated by a month. While the theory is that the clinicians learned what to watch and look for while evaluating in the training module, it is likely that it was not a frequent thought or activity throughout the interluding month and it is possible that attentiveness to certain details was overlooked.

Third, the evaluation process by the physician was not the same as that used for testing. In using the physician as the reference criterion, it was necessary for him to employ his own evaluation. This included an assessment of static posture. Though static posture alone did not make a decision about a participant’s rating, it was not something which was looked for in the SDT and therefore may have contributed to the high number

of dyskinetic ratings by the physician compared to the number of normal assessments by the clinicians.

Furthermore, this study relied on physician assessment as the reference criterion for diagnostic accuracy. Previous studies (Tate et al., 2009; Uhl et al., 2009) have used an objective method, three-dimensional motion analysis, to serve as the reference standard. Though data from Uhl et al. suggest that experts can be accurate in their assessments, they are less so than an objective measure. However, given the previous documentation regarding diagnostic accuracy of this SDT against three-dimensional analysis, the findings of this study do not contradict it. There may have been enhanced accuracy had three-dimensional analysis been the reference.

Fourth, the limited participant pool led to a high incidence of dyskinesia. The physician only rated one individual out of 17 to have bilaterally normal motion. As Fritz and Wainner (2001) noted, sensitivity and accuracy are significantly influenced in a negative way by high prevalence of a condition in the participant population. High prevalence, as evidenced here by a 94.4% occurrence of any dyskinesia in the participant pool and lack of variability between participants may explain the suppressed sensitivity and NPV. Certainly a higher number of participants with normal evaluations would have been desirable to provide a more balanced perspective.

Finally, in light of the high levels of agreement between raters previously reported, the choice of using Gwet's AC1 was made in order to ensure a robust reliability. By not utilizing a kappa statistic such as Fleiss', it may not be possible to adequately compare the reliability found in this study with those already published. It may also be of

value to go back and recalculate the reliability found in McClure et al. (2009), Uhl et al. (2009), and Dwelly (2012) utilizing the AC1 statistic to obtain more representative reliability values.

Theoretical application

The present study extends the body of literature in several ways. First, this is the first published study utilizing a rater pool that consisted only of certified athletic trainers. Previous studies have incorporated the occasional athletic trainer (McClure et al., 2009; Uhl et al., 2009) or student populations (Dwelly, 2012) but not one in which all of the raters were certified prior to testing.

This is important because it demonstrates the teachability of the method. Despite the relative lack of clinical experience of the current clinicians used, certified athletic trainers are expected to perform their job duties without the oversight and correction afforded by the student experience. Thus it is reasonable to presume that the certified athletic trainers in this study already possessed a preferred method for assessing dyskinesia. To be asked to shed that in deference to a very particular, hands-off method and to perform as reliably as well published experts in the field speaks to the value of the test.

Secondly, this is one of the first, if not the first, study in which no rating was done via video camera. By removing video-taped assessments from the equation, this ensured that the clinicians were viewing a three dimensional event rather than a three dimensional event delivered through a two dimensional medium. While it is not necessarily the case,

not being able to view participants live during the actual motions may have detracted from the ability of the raters to judge the motion because they were limited to a fixed view.

Viewing the participants live rather than on video recreates the clinical environment and experience. It is unlikely that an athletic trainer will have the luxury to have an athlete be video-taped for later, non-interactive evaluation. Rather, scapular motion screens are going to be part of the injury prevention or evaluation process for the shoulder complex. Having a test with such a high degree of inter-rater agreement and moderate inter-rater reliability enables the clinical athletic trainer to employ this test in the field knowing that there is a reasonable likelihood of generating similar results as an expert in the field.

Clinical Application

The results of this study continue to justify the incorporation of the yes/no SDT into both educational and clinical practice. Aside from a functional understanding of the normal motion of the scapula, this test requires no special skills and little need for the more complicated discrimination employed for the four-category scapular dyskinesis test (Kibler et al., 2002). The consistency of reliability across multiple clinician populations indicates its capability for widespread deployment. Its previously reported measures of diagnostic accuracy are not refuted by the present study and suggest that it is a sufficient screening tool for identifying or ruling out individuals with scapular dyskinesis. Variances between that data and the data found here can potentially be explained by the limitations previously noted.

By having a reliable, evidence-based tool for evaluating scapular motion, clinicians have a powerful method for assessing, preventing, and treating shoulder complex injuries that might have otherwise been prolonged by undiagnosed scapular pathomechanics. Dwelly (2012) noted that athletic trainers as inexperienced as seniors in college can correctly identify the muscles to be strengthened given a pathomechanical presentation. Thus both preventative training and post-injury rehabilitation are enhanced by having this test in the athletic trainer's evaluative arsenal.

Future Research Considerations

As previously noted, it may be valuable to return to the data from previous studies and examine it utilizing Gwet's AC1 statistic rather than a kappa-based statistic. Doing so may provide an enhanced view of the reliability of the yes/no SDT, particularly in the presence of high agreement.

The results of this study and those of Dwelly (2012) also suggest the need for three-dimensional analysis, similar to that of Uhl et al. (2009), when the clinicians are athletic trainers. Though three-dimensional methods are fraught with their own shortcomings when applied to scapular motion, their data are more objective as a reference than human evaluation.

Conclusion

The upper extremity screen is one of the most frequently used tools for the athletic trainer in the evaluation of shoulder injury and dysfunction. The ability to accurately

observe abnormal motion patterns is a key component of this screen. To this end, numerous tests have been identified to aid the clinician in the determination of normal versus abnormal motion relative to the motion of the scapula in relation to both the thorax and the humerus.

Most recently, the yes/no scapular dyskinesis test has been identified as having greater clinical value due in part both to its comparative simplicity over previous tests as well as in the diagnostic utility of the test as quantified by common statistical measures in the field of orthopedic examination. However, this has only been proven to be true for experienced clinicians in the field of shoulder and scapular mechanics. To date, no study had been published which examined the less experienced clinician.

This investigation involved the use of clinician assessment of scapular motion. The degree to which multiple raters agreed on an assessment and the accuracy with which it was used, even in light of limited experience, indicates that this test is appropriate for clinical practice and ought to be included in educational programs as part of the upper extremity evaluation process.

REFERENCES

- Alberg, A. J., Park, J. W., Hager, B. W., Brock, M. V., & Diener-West, M. (2004). The use of “overall accuracy” to evaluate the validity of screening or diagnostic tests. *Journal of General Internal Medicine*, 19, 460-465. doi:10.1111/j.1525-2004.30091.x.
- Burkhart, S. S., Morgan, C. D., & Kibler, W. B. (2003). The disabled throwing shoulder: Spectrum of pathology part III: The SICK scapula, scapular dyskinesis, the kinetic chain, and rehabilitation. *Arthroscopy*, 19, 641-661. doi:10.1016/S0749-8063(03)00389-X
- Cleland, J. (2005). *Orthopaedic clinical examination: An evidence-based approach for physical therapists*. Carlstadt, NJ: Icon Learning Systems.
- Cull, P. (Ed.). (1989). *The sourcebook of medical illustration*. Park Ridge, New Jersey: Parthenon Publishing Group.
- Dwelly, P. (2012). *Athletic training students' ability to identify scapular dyskinesis* (Doctoral dissertation, University of Arkansas). Retrieved from <http://proquest.umi.com/pqdlink?Ver=1&Exp=11-08-2016&FMT=7&DID=2407594971&RQT=309&attempt=1&cfc=1>
- Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, 61, 29-48. Retrieved from http://agreestat.com/research_papers/bjmisp2008_interrater.pdf
- Fritz, J. M., & Wainner, R. S. (2001). Examining diagnostic tests: an evidence-based

- perspective. *Physical Therapy*, 81, 1547-1564. Retrieved from <http://ptjournal.apta.org/content/81/9/1546.full.pdf+html>
- Kibler, W. B. (1998). The role of the scapula in athletic shoulder function. *The American Journal of Sports Medicine*, 26, 325-337. Retrieved from <http://ajsm.highwire.org/content/26/2/344.full.pdf+html>
- Kibler, W. B., Uhl, T. L., Maddux, J. W. Q., Brooks, P. V., Zeller, B., & McMullen, J. (2002). Qualitative clinical evaluation of scapular dysfunction: A reliability study. *Journal of Shoulder and Elbow Surgery*, 11, 550-556.
doi:10.1067/mse.2002.126766
- Ludewig, P. M., & Cook, T. M. (2000). Alterations in shoulder kinematics and associated muscle activity in people with symptoms of shoulder impingement. *Physical Therapy*, 80, 276-291. Retrieved from <http://www.ptjournal.apta.org>
- Lukasiewicz, A. C., McClure, P., Michener, L., Pratt, N., & Sennett, B. (1999). Comparison of 3-dimensional scapular position and orientation between subjects with and without shoulder impingement. *Journal of Orthopaedic and Sports Physical Therapy*, 29, 574-586. Retrieved from <http://www.jospt.org>
- McClure, P., Tate, A. R., Kareha, S., Irwin, D., & Zlupko, E. (2009). A clinical method for identifying scapular dyskinesis, Part 1: Reliability. *Journal of Athletic Training*, 44, 160-164. Retrieved from <http://www.journalofathletictraining.org>
- McQuade, K. J., Dawson, J., & Smidt, G. L. (1998). Scapulothoracic muscle fatigue associated with alterations in scapulohumeral rhythm kinematics during maximum resistive shoulder elevation. *Journal of Orthopaedic and Sports*

- Physical Therapy*, 28, 74-80. Retrieved from <http://www.jospt.org>
- McQuade, K. J., & Smidt, G. L. (1998). Dynamic scapulohumeral rhythm: The effects of external resistance during elevation of the arm in the scapular plane. *Journal of Orthopaedic and Sports Physical Therapy*, 27, 125-133. Retrieved from <http://www.jospt.org>
- Odom, C. J., Taylor, A. B., Hurd, C. E., & Denegar, C. (2001). Measurement of scapular asymmetry and assessment of shoulder dysfunction using lateral scapular slide test: A reliability and validity study. *Physical Therapy*, 81, 799-809.
- Schmitt, L., & Snyder-Mackler, L. (1999). Role of scapular stabilizers in etiology and treatment of impingement syndrome. *Journal of Orthopaedic and Sports Physical Therapy*, 29, 31-38. Retrieved from <http://www.jospt.org>
- Tate, A. R., McClure, P., Kareha, S., Irwin, D., & Barbe, M. F. (2009). A clinical method for identifying scapular dyskinesis, Part 2: Validity. *Journal of Athletic Training*, 44, 165-173. Retrieved from <http://www.journalofathletictraining.org>
- Uhl, T. L., Kibler, W. B., Gecewich, B., & Tripp, B. L. (2009). Evaluation of clinical assessment methods for scapular dyskinesis. *Arthroscopy*, 25, 1240-1248.
- Warner, J. J. P., Micheli, L. J., Arslanian, L. E., Kennedy, J., & Kennedy, R. (1992). Scapulothoracic motion in normal shoulders and shoulders with glenohumeral instability and impingement syndrome. *Clinical Orthopaedics and Related Research*, 285, 191-199. Retrieved from <http://ovidsp.tx.ovid.com>
- Warrens, M. J. (2010). A formal proof of a paradox associated with Cohen's kappa. *Journal of Classification*, 27, 322-332. Retrieved from

<http://www.sciweavers.org/publications/formal-proof-paradox-associated-cohens-kappa>

APPENDICES

Appendix A: Informed Consent

v6 2/3/2010



Department of HPER
7000 Old Main Hill
Logan UT 84322-7000
Telephone: (435) 797-1495



Page 1 of 2
USU IRB Approval: March 17, 2011
Amendment #1 Approved: 08/29/2011
Approval Terminates: 03/16/2012
Protocol #2884
IRB Password Protected per IRB Coordinator

INFORMED CONSENT

Reliability of the Yes/No Scapular Dyskinesis Test when used by Graduate Assistant Athletic Trainers

Introduction/ Purpose Professor Gerald Smith and graduate student Adam Raikes in the Department of Health, Physical Education, and Recreation at Utah State University are conducting a research study concerning the evaluation of scapular dyskinesis which is abnormal motion of the shoulder blade. You have been asked to take part either because you have been evaluated by Dr. Lyons to have scapular dyskinesis or because you were recruited from the general college populace.

Procedures If you agree to be in this research study, an evaluation of your shoulder motion involving the following procedures will occur:

1. *If you have not been previously evaluated by Dr. Lyons for scapular dyskinesis, a clinical exam by him will be performed.*
2. *You will be asked to either remove your shirt or wear clothing that will allow your entire shoulder blade to be seen.*
3. *You will be asked to lift a light dumbbell of 3 or 5 pounds.*
4. *You will be asked to perform 5 repetitions of shoulder flexion and 5 repetitions of scapular plane elevation while the motion of your shoulder blades is observed by graduate assistant athletic trainers.*

New Findings During the course of this research study, you will be informed of any significant new findings (either good or bad), such as changes in the risks or benefits resulting from participation in the research, or new alternatives to participation that might cause you to change your mind about continuing in the study. If new information is obtained that is relevant or useful to you, or if the procedures and/or methods change at any time throughout this study, your consent to continue participating in this study will be obtained again.

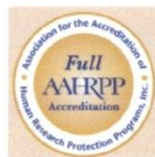
Risks Participation in this research study may involve some added risks or discomforts. For example, you may experience muscle soreness or strain consistent with a low-intensity weight-lifting workout.

Benefits There may or may not be any direct benefit to you from these procedures. The investigators, however, will learn more about the ability of athletic trainers to clinically evaluate scapular dyskinesis.

Explanation & offer to answer questions Adam Raikes has explained this research study to you and answered your questions. If you have other questions or research-related problems, you may reach Professor Smith at 797-8845.

Payment/Compensation You will not be receiving any compensation for your participation in this study. There is no cost to you for participation.

v6 2/3/2010



Page 2 of 2
USU IRB Approval: March 17, 2011
Amendment #1 Approved: 08/29/2011
Approval Terminates: 03/16/2012
Protocol #2884
IRB Password Protected per IRB Coordinator

INFORMED CONSENT

Reliability of the Yes/No Scapular Dyskinesis Test when used by Graduate Assistant Athletic Trainers

Voluntary nature of participation and right to withdraw without consequence Participation in research is entirely voluntary. You may refuse to participate or withdraw at any time without consequence or loss of benefits.

Confidentiality Research records will be kept confidential, consistent with federal and state regulations. Only Adam Raikes and Professor Smith will have access to the data which will be kept in a locked file cabinet in a locked room. Personal, identifiable information will be kept for a year while results are being analyzed and a thesis is being written but will be destroyed after that period. In written documents such as a thesis or scientific manuscripts, no personal identifiable information will be used.

IRB Approval Statement The Institutional Review Board for the protection of human participants at USU has approved this research study. If you have any pertinent questions or concerns about your rights or a research-related injury, you may contact the IRB Administrator at (435) 797-0567 or email irb@usu.edu. If you have a concern or complaint about the research and you would like to contact someone other than the research team, you may contact the IRB Administrator to obtain information or to offer input.

Copy of consent You have been given two copies of this Informed Consent. Please sign both copies and retain one copy for your files.

Investigator Statement "I certify that the research study has been explained to the individual, by me or my research staff, and that the individual understands the nature and purpose, the possible risks and benefits associated with taking part in this research study. Any questions that have been raised have been answered."

Signature of Principal Investigator and Student Researcher

Gerald Smith PhD
435-797-8845
Gerald.Smith@usu.edu

Adam Raikes
adam.raikes@gmail.com

Signature of Participant By signing below, I agree to participate.

Participant's signature

Date

Appendix B: Physician examination

1. Observation

- a. Gross symmetry in a static position
- b. Measure bilateral distance from medial scapula to spinal column
 - i. Arms hanging down
 1. Inferior angle
 2. Spine
 3. Superior angle
 - ii. Hands on hips
 1. Same three points
- c. Dynamic motion with the patient facing forward, looking from behind
 - i. Forward flexion
 - ii. Abduction
 - iii. Protraction
 - iv. Retraction
- d. Resisted Dynamic
 - i. Empty can test
 - ii. Resisted wall pushup
 - iii. Unresisted wall pushup

Determinations of scapular dyskinesis are based on the overall presentation during the entirety of the exam. The decision is made based on the persistence of dyskinesis throughout rather than intermittent, inconsistent presentation.

Appendix C: Clinician score sheet

Appendix D: Raw Data

Table D1

Clinician findings of normal or dyskinetic motion

Participant	Clinician									
	1	2	3	4	5	6	7	8	9	10
1	1	1	0	1	0	1	0	0	1	0
2	0	0	0	0	1	0	0	0	0	0
3	1	1	0	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1
6	1	1	0	1	1	1	1	1	1	0
7	1	1	0	0	1	0	1	0	1	0
8	1	1	1	1	1	1	1	1	1	1
9		0	1	0	1	1	1	0	0	1
10	1		0	1	1	1	1	0	0	0
11	1	0		1	0	1	0	0	1	0
12	0	1	0		0	1	1	0	0	1
13	0	0	0	0		0	0	0	0	0
14	1	1	0	0	0		1	0	0	0
15	1	1	1	1	1	1		1	1	1
16	1	1	0	1	1	1	1		1	1
17	1	1	1	1	1	1	1	1		1

Note. 0 = Normal motion observed; 1 = Dyskinetic motion observed

Table D2

Clinician findings of normal or dyskinetic motion per side

Participant	Clinician									
	1	2	3	4	5	6	7	8	9	10
1	1	3	0	3	0	3	0	0	3	0
2	0	0	0	0	1	0	0	0	0	0
3	2	2	0	2	2	2	2	2	2	2
4	3	3	3	3	3	3	3	3	3	3
5	1	1	1	1	3	1	3	1	3	1
6	3	3	0	2	3	1	3	2	3	0
7	3	3	0	0	3	0	1	0	3	0
8	3	3	2	1	3	3	3	3	3	4
9		0	4	0	2	2	2	0	0	2
10	2		0	2	2	2	3	0	0	0
11	3	0		1	0	1	0	0	3	0
12	0	3	0		0	2	2	0	0	4
13	0	0	0	0		0	0	0	0	0
14	2	3	0	0	0		2	0	0	0
15	3	1	1	3	3	3		1	2	1
16	3	3	0	3	3	3	3		3	4
17	3	3	1	3	3	3	3	3		3

Note. 0 = Normal motion observed; 1 = Dyskinesia observed on the left side; 2 = Dyskinesia observed on the right side 3 = Dyskinesia observed on both sides; 4 = Dyskinesia observed, uncertain of side

Table D3

Physician and clinician findings of normal or dyskinetic motion isolated to the left scapula

Participant	Clinician										
	MD	1	2	3	4	5	6	7	8	9	10
1	0	1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	1	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0
4	1	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	0	1	0	1	0	1
6	1	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	0	0	0	1	0	0	0
8	0	0	0	0	1	0	0	0	0	0	0
9	0		0	0	0	0	0	0	0	0	0
10	0	0		0	0	0	0	0	0	0	0
11	1	0	0		1	0	1	0	0	0	0
12	0	0	0	0		0	0	0	0	0	0
13	0	0	0	0	0		0	0	0	0	0
14	0	0	0	0	0	0		0	0	0	0
15	0	0	1	1	0	0	0		1	0	1
16	0	0	0	0	0	0	0	0		0	0
17	0	0	0	1	0	0	0	0	0		0

Note. 0 = Normal motion observed on the left side; 1 = Dyskinetic motion observed on the left side

Table D4

Physician and clinician findings of normal or dyskinetic motion isolated to the right scapula

Participant	Clinician										
	MD	1	2	3	4	5	6	7	8	9	10
1	1	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0
3	1	1	1	0	1	1	1	1	1	1	1
4	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	1	0	0	0	1	0	0
7	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0	0	0	0
9	0		0	0	0	1	1	1	0	0	1
10	1	1		0	1	1	1	0	0	0	0
11	0	0	0		0	0	0	0	0	0	0
12	0	0	0	0		0	1	1	0	0	0
13	1	0	0	0	0		0	0	0	0	0
14	0	1	0	0	0	0		1	0	0	0
15	0	0	0	0	0	0	0		0	0	0
16	0	0	0	0	0	0	0	0		0	0
17	0	0	0	0	0	0	0	0	0		0

Note. 0 = Normal motion observed on the right side; 1 = Dyskinetic motion observed on the right side

Table D5

Physician and clinician findings of normal or dyskinetic motion in both scapulae

Participant	Clinician										
	MD	1	2	3	4	5	6	7	8	9	10
1	0	0	1	0	1	0	1	0	0	1	0
2	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0
4	0	1	1	1	1	1	1	1	1	1	1
5	0	0	0	0	0	1	0	1	0	1	0
6	0	1	1	0	0	1	0	1	0	1	0
7	0	1	1	0	0	1	0	0	0	1	0
8	1	1	1	0	0	1	1	1	1	1	0
9	1		0	0	0	0	0	0	0	0	0
10	0	0		0	0	0	0	1	0	0	0
11	0	1	0		0	0	0	0	0	1	0
12	1	0	1	0		0	0	0	0	0	0
13	0	0	0	0	0		0	0	0	0	0
14	1	0	1	0	0	0		0	0	0	0
15	1	1	0	0	1	1	1		0	1	0
16	1	1	1	0	1	1	1	1		1	0
17	1	1	1	0	1	1	1	1	1		1

Note. 0 = Bilaterally normal motion observed; 1 = Bilaterally dyskinetic motion observed

Table D6

Physician and clinician findings of any dyskinesia

Participant	Clinician										
	MD	1	2	3	4	5	6	7	8	9	10
1	1	1	1	0	1	0	1	0	0	1	0
2	1	0	0	0	0	1	0	0	0	0	0
3	1	1	1	0	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	0	1	1	1	1	1	1	0
7	0	1	1	0	0	1	0	1	0	1	0
8	1	1	1	1	1	1	1	1	1	1	1
9	1		0	1	0	1	1	1	0	0	1
10	1	1		0	1	1	1	1	0	0	0
11	1	1	0		1	0	1	0	0	1	0
12	1	0	1	0		0	1	1	0	0	1
13	1	0	0	0	0		0	0	0	0	0
14	1	1	1	0	0	0		1	0	0	0
15	1	1	1	1	1	1	1		1	1	1
16	1	1	1	0	1	1	1	1		1	1
17	1	1	1	1	1	1	1	1	1		1

Note. 0 = Bilaterally normal motion observed; 1 = Dyskinetic motion observed